Jurnal Sains Sukan dan Pendidikan Jasmani Vol 11, No 1, 2022 (45-57)
ISSN: 2232-1918 / eISSN: 2600-9323
https://ejournal.upsi.edu.my/journal/JSSPJ

Validity and Reliability of Handball Cognitive Assessment Instrument for Year 4: Use of Rasch Model

Uthaya Chanthira Kumar AL Muniandy[1], Hutkemri Zulnaidi[1], Siti Hajar Halili[1]

[1]Faculty of Education, University of Malaya (UM), Kuala Lumpur, MALAYSIA

Published online: 20 June 2022

To cite this article (APA): Muniandy, U. C. K., Zulnaidi, H., & Halili, S. H. (2022). Validity and Reliability of Handball Cognitive Assessment Instrument for Year 4: Use of Rasch Model. *Jurnal Sains Sukan & Pendidikan Jasmani, 11*(1), 45-57. https://doi.org/10.37134/jsspj.vol11.1.6.2022

To link to this article: https://doi.org/10.37134/jsspj.vol11.1.6.2022

ABSTRACT

The Handball Cognitive Assessment instrument for the Year 4 Physical Education (PE) subject is developed with the aim to assess the capabilities of the cognitive domain for the category of handball attack-play in regard to knowledge and understanding. In ensuring the instrument developed is dependable and suitable for use, this study was conducted to analyze and investigate empirical evidence on the validity and reliability of the instrument by using the Rasch analysis model. The Handball Cognitive Assessment instrument consists of twenty-five items that have been piloted by 51 Year 5 primary school students. The WINSTEPS programme was utilized to conduct the Rasch model analysis to obtain the validity and reliability of the instrument. The test of unidimensionality for the twenty-five items has a measured variance of 25.4% and all items exceeded the minimum of 20.0% as required by the Rasch model. The reliability index of the items is 0.83 and the reliability index of the respondents is 0.54. All the items have a positive correlation with PMC (PT-MEASURE CORR) and the item polarity shows that there is no discrepancy between the items for the constructs measured. The Outfit-MNSQ value shows a value that is less or equal to 1.5 logit which is defined as productive. However, for items 6, 8, and 14, although the Outfit-MNSQ value exceeds 1.5, it did not diverge too far when compared to the pilot study. These items were scrutinized again, and the validity is revalidated. The findings of this study prove that the Handball Cognitive Assessment is able to be utilized as an assessment instrument for the cognitive domain of PE specifically the Handball activity for Year 4 primary school students as it has proper validity and reliability that is accurate and can be implemented empirically.

Keywords: Validity, reliability, cognitive, instrument, Rasch model

INTRODUCTION

The subject of Physical Education (P.E.) is a unique education curriculum as it consists of learning and evaluating from three central domains which are psychomotor, cognitive, and affective (Liza et al.,2016). Assessment of learning achievement from these domains at the optimum level can be achieved with the implementation of holistic assessment (Jani et al., 2014). Therefore, the Curriculum Development Board has rebranded the School-Based Assessment to Class Based Assessment in which teachers play an important role in determining the learning objective and improving student learning through quality assessment that is categorized as assessment for learning, assessment as learning, and assessment of learning (Ministry of Education Malaysia, 2018). Thus, these three domains are interrelated with each other,

where learning involves the psychomotor domain and requires students to have a background theory of the cognitive and affective domains (Thoirs & Coffee, 2012). Therefore, teaching and learning concerning the cognitive aspect is important in PE because students will be more inclined to understand the rules and strategy development of the movement in learning as well as improve their motor skills in doing physical sports activity alongside improving thinking capabilities and interpretation skills (Anderson et al., 2001).

Zhang et al. (2019) in their study stated that the cognitive domain and knowledge is a dimension that refers to the understanding of a student regarding a topic whether through factual knowledge, conceptual knowledge, or procedural knowledge. Between these three, PE involves factual knowledge such as rules and principles of motor skills and procedural knowledge such as demonstration and operation of motor skills that contributes to metacognitive development (Chang et al., 2020). Thus, the cognitive domain plays a vital role in the learning of sports activities in PE and needs to be given more attention and focus.

The cognitive domain is defined as the intellectual and thinking ability in which cognitive assessments are often utilized to evaluate students' knowledge and intelligence (Kamarudin & Halili, 2004) which can transpires at any moment and in any learning context (Bhasah & Bakar, 2008). Bloom et al., (1956) suggested six cognitive levels according to the original taxonomy which consists of fact memorization at the lowest level, followed by mental processes that are more complex and abstract (synthesize, analyze), and finally at the highest level is evaluation skills (Adams, 2015). The initial level of the cognitive domain is linked to the process of recalling material that has been learned. The recalling process is the foundation for cognitive tendency, which brings about a creative combination of ideas to synthesize them as something new.

Although the original taxonomy is in frequent use in the education sector, it did receive strong criticism due to its nature of linear hierarchy from simple to complex and the way that the taxonomy characterizes students' knowledge acceptance (Amer, 2006). Additionally, Bloom's taxonomy is found to not have direct practical usage in curriculum development (Cannon & Feinstein, 2005) and has a strong dependency towards actions (Amer, 2006; Ferris & Aziz, 2005).

As a result of the criticism that arose, the taxonomy was reviewed and revised by a group of cognitive psychologists, curriculum researchers, and assessment and testing experts which was led by Anderson in 2001 to reflect its relevancy in 21st-century teaching and learning (Amer, 2006; Anderson et al., 2001; Cannon & Feinstein, 2005). The changes in Bloom's Taxonomy are shown in Figure 1 below.

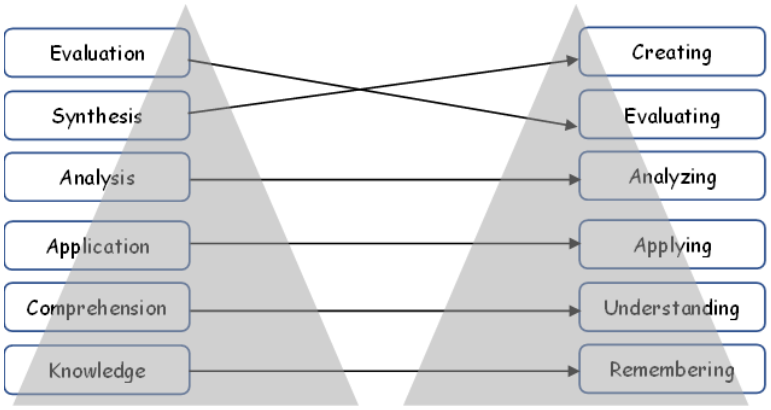

Figure 1. Bloom's Taxonomy (1956) and Anderson's Revised Bloom's Taxonomy (2001)

The revised taxonomy has undergone a few significant changes in regards of assumptions, taxonomy structure, and use of terms. The most momentous change that was observed is the change of noun (dimension of knowledge) to verb (dimension of remembering) of the components of Bloom's taxonomy. The usage of verbs is more accurate as the taxonomy defines the various thinking and it is an active process. The level of synthesis and evaluation has been rearranged, and all components had been replaced with

verbs. Although the need to evaluate and assess learning outcomes of PE has been extensively discussed in past studies, there is a lack of studies conducted on the dimensions of subcategories of knowledge and understanding (Ayers, 2004).

This has to be stressed because the holistic development of a student through knowledge and experience that involves the cognitive domain is the key element in the learning of PE as it contributes to successful learning of physical activities (Anderson et al., 2001; Darst & Pangrazi, 2009; Chang et al., 2020). Therefore, teachers need to prioritize and emphasize the cognitive learning domain by enhancing focus and developing student learning so that it can increase knowledge and understanding about the topic which then can be applied in the sports activity (Gurvitch & Metzler, 2013). According to Nixon and Locke (1973), the emphasis of the cognitive learning domain not a new issue and students need to generate their personal cognitive image of specific movement and learning through imitation of behavior which can be done with existing knowledge and understanding. This is in line with Piaget's (1973) theory of constructivism which focuses on experience and observation from the environment as a stimulus that is experienced by the student during learning which will then assimilate into the existing cognitive process. Therefore, the imitation process will give meaning to the student self when the behavior can be conducted autonomously without any extra encouragement and based on existing knowledge. However, despite its importance, it is found that teachers do not pay adequate attention to developing sports games that put focus on the cognitive domain (Balakrishnan et al., 2011).

Through an interview with PE teachers, Liza et al. (2016) found that the domains of affective and cognitive are usually assessed through summative examinations that are usually conducted during the middle or end of the year while the psychomotor domain is assessed through observation during teaching and learning. This study is comparable to Hassan and Abdullah, (2016), who discovered that teachers exclusively undertake summative evaluations and are perceived to lack competence in assessment management. This problem arose as a result of most PE teachers are not well trained and adept in developing assessment items. Likewise, schoolteachers that conduct assessments did not appropriately evaluate PE and there is no standardized document in use by PE teachers to evaluate the student achievement level and understanding of the topics/activities being taught (Khamis et al., 2018). This is further strengthened by the study of needs analysis by Gengatharan and Azli Rahmat (2019) who found that 46.7% of teachers did not conduct an assessment after completing teaching and learning, despite the component being readily available as supporting teaching material (reference: *SPI Bil* 14, 2018) by the Ministry of Education (MoE) Malaysia. The majority of teachers are found to be utilizing third-party reference books sold at bookstores as the main reference material to build instruments or download instruments from ready-made resources on the internet and even some teachers use self-designed questions based on their understanding. Time constraints and teachers 'commitment to various tasks in school cause teachers to find devise simple ways in preparing questions for assessment purposes. Abdullah et al. (2015) addressed this problem in their findings, which revealed that 66.7% of instructors did not construct any form of instrument owing to their excessive task commitment and time limits. Even worse, teachers at schools face a lack the competence to build their own items. As per finding by Othman et al. (2013) found that most teachers who teach PE are still lacking competency in developing the required assessment items, lack understanding of assessments, are unfamiliar with the concept of validity and reliability in building an assessment item, as well as unskilled in executing and evaluating the said assessment.

From the aspect of learning, students' understanding plays a significant role in facing new learning situations. When discrepancies exist in their understanding and learning situation, students will try to enhance their cognitive knowledge to accommodate it. Thus, during a sports activity, students can respond to unexpected situations that may not be practiced during their training sessions (Balakrishnan et al., 2011). Therefore, based on the issues that has been put forward, a worrying situation arises as there is indeed a large discrepancy in the cognitive domain assessment in teaching and learning based on Bloom's taxonomy in which cognitive skills are more challenging, contribute more to meaningful learning and requires more attention and training (Mohan, 2018). As such, this study aims to develop an instrument for cognitive

assessment for the learning of Handball for Year 4 primary school students as well as ensure the instrument's validity and reliability.

METHODOLOGY

Study Sample

A total of fifty-one students (29 male and 22 female) aged 11 years old that took PE class and learned Handball in primary school were chosen as a sample for this study. The intact sampling method was utilized to randomly select the sample from the list of schools.

Study Instrument

Cognitive learning assessment of Handball for the subject of Physical Education is done by the pencil and paper testing method. For this study, the pencil and paper testing method refers to objective question items which is the traditional testing method utilized to observe students' achievements (Akın, 2022). A quality test is a set of instruments that is capable of measuring what needs to be measured and is suitable for the target audience. The Handball Cognitive Assessment Instrument was created specifically for this study.

Table 1: Table of Specifications Handball Cognitive Assessment Instrument

Skill	Domain	No. Of Questions Difficulty Level						Quantity
		Remembering	Understanding	Applying	Analyzing	Evaluating	Creating	
Throwing/ Catching	Cognitif			2,5,6,21				4
	Psychomotor	1,4,24	3					4
	Affective							-
Dribbling	Cognitive			8	9,19,20			4
	Psychomotor			7				1
	Affective							-
Blocking	Cognitive	10,11			12			3
	Psychomotor							-
	Affective							-
Intercepting	Cognitive		15	14	13			3
	Psychomotor			23			22	2
	Affective							-
Tackling/ Shooting	Cognitive							-
	Psychomotor	16				17,18		3
	Affective						25	1
Total No. Of Questions		6	2	8	5	2	2	25
Weightage (%)		25	10	35	20	5	5	100%
Category of Difficulty		**Lower Order Thinking Skills (LOTS)**			**Higher Order Thinking Skills (HOTS)**			

Source: Adapted from Lacy, A. C., & Hastad, D. N. (2003). *Measurement and Evaluation in Physical Education and Exercise Science* (4th ed.). Pearson.

The pencil and paper test in this study consists of 25 questions that follow the Cognitive domain in accordance with Anderson's Revised Bloom's Taxonomy (2001) which are (6) six questions of Remembering (25.00%), (2) two questions of Understanding (10.00%), (8) eight questions of Applying (35.00%), (5) five questions of Analyzing (20.00%), (2) two questions of Evaluating (5.00%) and (2) two questions of Creating (5.00%). Table of Specifications Handball Cognitive Assessment Instrument and the rubric of questions are all related to the game's skill, technique, and strategy based on Lacy and Hastad, (2003); Lacy and Williams, (2018). The objective question of this instrument consists of four choices of answers (A, B, C, and D) and only one choice is correct out of the four available. The scoring for Cognitive learning is based on Anderson's Revised Bloom's Taxonomy (2001) which applies the testing method of pencil and paper utilizes the scale of achievement as shown in Table 2 below.

Table 2: Scale and Achievement Level of the Handball Cognitive Assessment Instrument

Scale	Level
80 to 100	Excellent
60 to 79	Good
40 to 59	Satisfactory
20 to 39	Weak
19 to 0	Very Weak

Study Procedure

Criteria validity and content validity are the two main validity tests that were conducted during the development of the instrument. The instrument went through a rigorous process of validity checking that was done twice. At the initial stage, the researcher has verified the instrument with experts from the aspect of criteria and instrument content from two expert panels in PE, one senior lecturer of PE, one handball state coach, and one linguist. Firstly, language validity is done before the instrument was given to the content experts to ensure that the language used is accurate and precise. After the instrument was given feedback and approval by the content experts, the researcher manually conducted an approval statistic to find out the content validity index from the panel of experts based on the following formula:

$$\frac{\text{Total Score of Experts}}{\text{Maximum Score of Experts}} = \text{Content Validity Index (CVI)}$$

Source: Ahmad (2014) page 74

In which.
Total Score of Experts = sum of all experts that were approved according to the scale
Maksimum skor of Experts = Number of Items x Number of Scale

The approval of content experts is as shown in Table 3 in which the content validity of the Handball Cognitive Assessment Instrument is CVI = 0.82, (N=5) for the cognitive domain in Primary Year 4 handball.

Table 3: Initial Content Validity of Handball Cognitive Assessment Instrument (N=5)

Element	Item	Panel of Experts					∑	M
		Language Expert	PE Expert	PE Expert	PE Lecturer	Handball Coach		
Cognitive Domain: *Handball Cognitive Assessment Instrument*	Questions are accurate with learning content	3	3	3	3	3		
	Text of the questions is suitable and precise	3	3	3	4	4		
	Photo/Figure/Table is accurate and suitable	4	3	3	4	4		
	Questions follow the hierarchy of the revised Bloom's Taxonomy (2001)	3	3	2	3	3		
	Questions are in accordance with age level and understanding	3	3	3	3	3		

Scale and scoring are suitable		4	4	4	3	4	
	∑	20	19	18	20	21	102
	M	0.83	0.79	0.75	0.83	0.88	**0.82**

Next, the researcher has improved the Handball Cognitive Assessment Instrument based on the comments and suggestions made by the panel of experts, linguists, and coaches based on the curriculum needs of Year 4 PE. After the improvement was meticulously done, the researcher has conducted the validation process for the second time which involved a panel of six experts which consists of four experts of PE who are Ph.D. holders, one linguist with a Ph.D. holder from the Department of Language and Literacy, and one state coach that was referred to obtain content validity for the items listed. The approval of content experts is as shown in Table 4 in which the second round of content validity of the Handball Cognitive Assessment Instrument is CVI = 0.93, (N=6) for the cognitive domain in Primary Year 4 handball. Therefore, the validity by experts for Handball Cognitive Assessment Instrument is accepted based on Polit et al., (2007) who stated that the index value of content validity needs to be at a minimum of 0.83.

Table 4: Second Content Validity of the Handball Cognitive Assessment Instrument by Experts (N=6)

Element	**Item**	**Panel of Experts**							
		Language Expert	**PE Expert**	**PE Expert**	**PE Lecturer**	**Handball Coach**	**Sport Lecturer**	∑	**M**
Cognitive Domain: *Handball Cognitive Assessment Instrument*	Questions are accurate with learning content	4	4	3	3	4	4		
	Text of questions is suitable and precise	3	4	3	4	4	4		
	Photo/Figure/Table is accurate and suitable	4	4	4	4	4	4		
	Questions follow the hierarchy of the revised Bloom's Taxonomy (2001)	4	3	3	4	4	3		
	Questions are in accordance with age level and understanding	4	3	3	4	4	4		
	Scale and scoring are suitable	4	4	4	4	4	4		
	∑	23	22	20	23	24	23	112	
	M	0.96	0.92	0.83	0.96	1.00	0.96	**0.93**	

The panel of experts for both first and second validity is different individuals. This is to ensure there is no bias from the panels and a solid improvement can be done to the instrument. After the validity is obtained and verified by the experts, Handball Cognitive Assessment Instrument was administered using the method of test and retest. At first, students that were chosen went through teaching and learning of Handball in PE for 4 weeks via self-learning in which Year 5 students learned subchapters for Handball via online learning by using videos that were supervised by the PE teacher. After the learning period is

completed, the students were given the instrument to be answered based on their knowledge and understanding of the topics of Handball that were learned. The administration of the Paper-Pencil test for the instrument was conducted via a quiz in Google Forms. The answers were obtained, and marks were recorded. Next, the same student was retested after a period of two weeks using the same instrument, but the order of the items was shuffled, and this was done to avoid the influence of the pre-test that the students had taken two weeks earlier. The answers obtained and marks for the post-test were recorded. The evaluation of items is based on the answers of true or false, if the answer is true then the value (1) is given, and if false, a null value (0) is given.

Study Analysis

All the data that were obtained in this study were analyzed using WINSTEPS version 3.68.2 (Linacre, 2009) to obtain the validity analysis of the results and reliability of items using the Rasch Model. Analysis done via the Rasch Model generated statistical analysis fit that gives the information to the researcher on whether the data obtained is ideal and shows the students' capabilities based on the answer patterns according to the difficulty level of the items (Misbach & Sumintono, 2014). Via the Rasch model, the validity and reliability of the instrument can be determined by looking at various analyses such as item polarity, unidimensionality, item-individual/respondent mapping, item-individual reliability, and some other forms of analysis (Bond & Fox, 2015). Therefore, this study was conducted to obtain empirical evidence related to the validity and reliability of items developed to measure the knowledge and understanding based on cognitive domains.

RESULTS

Construct Validity for Handball Cognitive Assessment Instrument

Construct validity was done using the Rasch analysis for dichotomy data. According to Nurfaizin (2019), construct validity is a type of validity that places importance on how far the item measure can what needs to be measured following the concepts and definitions that were set. Therefore, construct validity is conducted after obtaining criteria and content validity from the experts. Analysis using the Rasch Model for the cognitive domain was done on the items to measure the construct validity of the Handball Cognitive Assessment Instrument. According to Figure 2 and Figure 3, it is found that all question items have a positive Point Measurement Correlation (PT-MEASURE CORR). This shows that there is no conflict between the item and the measurement of the questions. Additionally, if we look at the value of Outfit-MNSQ, it can be deduced that almost all items have a logit value that is less than 1.5, and only three items that exceed the Outfit-MNSQ of 1.5 which are item 8 (1.64), item 14 (1.56) and item 6 (2.05). Item 8 obtained Outfit-ZSTD 0.5 while item 6 obtained Outfit-ZSTD 2.3 and item 8 obtained Outfit-ZSTD 0.7. Compared to the Test 1 phase data, the value of Outfit-MNSQ for item 8 (1.65), item 14 (1.50), and item 6 (1.77), while item 8 obtained Outfit-ZSTD 0.1. Furthermore, item 6 obtained Outfit-ZSTD 2.2. and item 8 obtained Outfit-ZSTD 0.7. As Test 1 showed a value that is nearest to the minimum logit value of 1.5, the researchers concluded that it is acceptable to maintain all the items of the cognitive instrument (Erfan et al., 2020).

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
22	20	49	2.28	.32	.89	-.9	.84	-.9	.51	.42	72.9	68.3	S22
18	23	49	1.98	.31	.99	-.1	.95	-.2	.43	.41	66.7	65.9	S18
12	25	49	1.78	.31	.98	-.1	.90	-.6	.44	.40	62.5	65.2	S12
17	31	49	1.18	.32	1.09	.7	1.27	1.4	.28	.38	66.7	68.7	S17
23	31	49	1.18	.32	.89	-.9	.82	-1.0	.49	.38	70.8	68.7	S23
13	33	49	.97	.33	1.20	1.4	1.10	.5	.23	.37	60.4	71.2	S13
8	34	49	.86	.34	1.37	2.4	1.65	2.4	.01	.36	58.3	72.5	S8
25	36	49	.62	.35	.89	-.6	.77	-.8	.46	.35	75.0	75.5	S25
2	37	49	.50	.36	1.25	1.3	1.37	1.2	.12	.34	70.8	77.1	S2
20	37	49	.50	.36	1.00	.0	1.04	.2	.34	.34	79.2	77.1	S20
21	37	49	.50	.36	.92	-.4	.77	-.7	.44	.34	79.2	77.1	S21
14	38	49	.37	.37	1.11	.6	1.50	1.5	.22	.33	77.1	78.7	S14
5	39	49	.23	.38	1.00	.1	1.07	.3	.32	.32	79.2	80.5	S5
10	40	49	.08	.39	.80	-.9	.62	-1.0	.51	.31	85.4	82.3	S10
3	41	49	-.09	.41	1.17	.7	1.07	.3	.18	.30	79.2	84.1	S3
24	41	49	-.09	.41	.86	-.5	1.00	.2	.40	.30	87.5	84.1	S24
9	42	49	-.26	.43	1.05	.3	1.32	.8	.22	.29	85.4	85.8	S9
19	43	49	-.46	.46	.95	.0	.70	-.5	.35	.27	89.6	87.7	S19
6	45	49	-.96	.55	1.16	.5	1.77	1.1	.07	.23	91.7	91.8	S6
11	45	49	-.96	.55	.77	-.5	.48	-.7	.45	.23	91.7	91.8	S11
15	45	49	-.96	.55	.89	-.1	.54	-.6	.38	.23	91.7	91.8	S15
4	47	49	-1.75	.75	.88	.0	.30	-.6	.36	.17	95.8	95.9	S4
1	48	49	-2.50	1.04	.92	.2	.21	-.4	.30	.12	97.9	98.0	S1
7	48	49	-2.50	1.04	1.01	.3	.39	-.1	.21	.12	97.9	98.0	S7
16	48	49	-2.50	1.04	1.06	.4	.74	.2	.11	.12	97.9	98.0	S16
MEAN	38.2	49.0	.00	.48	1.00	.2	.93	.1			80.4	81.4	
S.D.	7.6	.0	1.31	.23	.14	.8	.40	.9			12.0	10.4	

Figure 2: Item Polarity of Test 1 Administered

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
22	21	49	2.29	.32	.91	-.7	.86	-.8	.52	.44	72.3	67.9	S22
18	23	49	2.08	.32	.95	-.4	.93	-.4	.48	.44	70.2	66.2	S18
12	25	49	1.88	.32	.98	-.1	.90	-.6	.46	.43	57.4	65.3	S12
17	31	49	1.26	.33	1.08	.7	1.24	1.3	.32	.40	66.0	68.8	S17
23	31	49	1.26	.33	.90	-.8	.83	-.9	.49	.40	70.2	68.8	S23
8	33	49	1.05	.33	1.40	2.6	1.64	2.6	.05	.39	55.3	71.0	S8
13	34	49	.93	.34	1.21	1.4	1.11	.5	.24	.38	61.7	72.4	S13
21	37	49	.57	.36	.91	-.4	.76	-.8	.45	.36	78.7	76.9	S21
25	37	49	.57	.36	.89	-.6	.77	-.8	.46	.36	78.7	76.9	S25
2	38	49	.43	.37	1.26	1.3	1.42	1.3	.13	.35	72.3	78.5	S2
20	38	49	.43	.37	1.00	.1	1.06	.3	.34	.35	80.9	78.5	S20
14	39	49	.29	.38	1.10	.5	1.56	1.5	.23	.34	78.7	80.3	S14
5	40	49	.14	.40	.99	.0	1.08	.3	.33	.32	80.9	82.1	S5
10	40	49	.14	.40	.79	-.9	.61	-1.1	.51	.32	85.1	82.1	S10
3	41	49	-.02	.41	1.18	.8	1.06	.3	.20	.31	78.7	83.9	S3
24	41	49	-.02	.41	.86	-.5	.99	.1	.40	.31	87.2	83.9	S24
9	42	49	-.20	.44	1.06	.3	1.30	.7	.23	.30	85.1	85.6	S9
19	44	49	-.64	.50	.94	-.1	.67	-.4	.34	.26	89.4	89.5	S19
11	45	49	-.91	.55	.76	-.5	.47	-.7	.44	.24	91.5	91.6	S11
6	46	49	-1.25	.62	1.14	.4	2.05	1.3	.07	.21	93.6	93.7	S6
15	46	49	-1.25	.62	.86	-.1	.42	-.6	.37	.21	93.6	93.7	S15
4	47	49	-1.70	.75	.87	.0	.30	-.6	.35	.18	95.7	95.8	S4
1	48	49	-2.45	1.04	.91	.2	.21	-.4	.28	.13	97.9	97.9	S1
7	48	49	-2.45	1.04	1.01	.3	.38	-.1	.20	.13	97.9	97.9	S7
16	48	49	-2.45	1.04	1.07	.4	.73	.2	.12	.13	97.9	97.9	S16
MEAN	38.5	49.0	.00	.49	1.00	.2	.93	.1			80.7	81.9	
S.D.	7.6	.0	1.34	.23	.15	.8	.43	.9			12.4	10.5	

Figure 3: Item Polarity of Test 2 (repeat test) Administered

The next analysis conducted is to find out the unidimensionality of items by referring to the Principal Component Variance Analysis as shown in Figure 4. The test of unidimensionality is a form of test that needs to be conducted to find out the validity of the instrument (Andrich, 1988). Figure 4 below shows the total of items that were observed which are 25 items and has a measured variance of 25.4%, and all items exceeded the minimum requirement of 20% (Sumintono & Widhiarso, 2015). According to

Ramdani et al., (2020), if the measures can explain the raw variation to a ≥ 20 %, it shows there is an argument can be made for unidimensional measurement. The following are the interpretation criteria: sufficient if the percentage is between 20% and 40%, good if the percentage is between 40% and 60%, and incredibly good if the percentage is over 60%. Additionally, the unexplained variance value for the first contra is less than 8.4%, not exceeding the control limit of 15% (Sumintono & Widhiarso, 2014).

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		-- Empirical --		Modeled
Total raw variance in observations	=	33.5	100.0%	100.0%
Raw variance explained by measures	=	8.5	25.4%	26.7%
Raw variance explained by persons	=	3.7	11.0%	11.6%
Raw Variance explained by items	=	4.8	14.4%	15.1%
Raw unexplained variance (total)	=	25.0	74.6% 100.0%	73.3%
Unexplned variance in 1st contrast	=	2.8	8.4% 11.2%	
Unexplned variance in 2nd contrast	=	2.7	8.0% 10.7%	
Unexplned variance in 3rd contrast	=	2.5	7.4% 9.9%	
Unexplned variance in 4th contrast	=	2.2	6.6% 8.9%	
Unexplned variance in 5th contrast	=	1.9	5.6% 7.5%	

Figure 4: Principal Component Analysis (PCA)

Reliability of Handball Cognitive Assessment Instrument

Fahruna and Fahmi (2017) stated that reliability refers to the degree of measurement of the instrument that is used in the study to obtain the required information and can be relied on as the tool of data collection as well as expose the true conditions and situation on the ground. The Handball Cognitive Assessment Instrument that was developed by the researcher with reference from experts in PE and experienced PE teachers that have taught PE for more than 10 years is aimed to evaluate the level of knowledge and understanding skills in learning Handball. The aspects that were inspected in the developed instrument are its usage and language suitability with the age group and learning in school, the use of specific sentences and comprehensibility that are easy to understand, questions constructed to meet the intended purpose to be tested and obtained, the rubric of questions build according to learning syllabus and Bloom's Taxonomy based on Lacy and Hastad (2002) and lastly scoring based on Anderson's Revised Bloom's Taxonomy and Krathwohl (2001). The developed instrument that was built with reference from experts was discussed with teachers and a piloting test was conducted to find out if changes and improvements were needed. The procedure to verify the reliability of the instrument was done using analyses by Rasch Model for both individual reliability and items while the individual isolation index shows a strata amount of capability in a group of a sample while the item isolation index gives information on how far the sample respondents can show the isolation of item difficulty (Boone & Noltemeyer, 2017).

Table 5: Reliability Index and Isolation Index

Tests Conducted		Reliability Index	Isolation Index
Test 1 (Test)	Individual	0.54	1.07
	Item	0.83	2.19
	Value Alpha Cronbach (KR-20)	0.65	
Test 2 (Repeat Test)	Individual	0.54	1.09
	Item	0.83	2.25
	Value Alpha Cronbach (KR-20)	0.68	

Table 5 shows the summary of statistic output and the analysis of item/question and individual respondents obtained for the repeated test that was done after a two-week period from the primary test. The reliability of the items can be seen in the reliability of the test 1 and the repeated test 2 is the same in value, which is 0.83, which the magnitude of this reliability is considered sufficient (Sumintono & Widhiarso, 2015). Furthermore, the isolation index value of 2.19 and 2.25 was also obtained, in which an isolation index value that exceeds the value of 2.0 is able to differentiate the items of the instrument and difficulty items. The larger the value of the isolation index, the overall quality of the instrument is the higher from the sense of respondents and items as it can identify the target respondents and grouping of items (Sumintono & Widhiarso, 2015; Erfan et al., 2020).

As for individual reliability, the value is 0.54 in both the primary test and repeated test in which the value is categorized as weak by Sumintono and Widhiarso, (2015). The isolation index also did not achieve the range of 2.0 as both obtained a value of only 1.07 and 1.09. It can be concluded that the consistency of answers from the respondents is low, but the quality of the items of the instruments is good.

Additionally, the number of respondents that were only 49 Year 5 pupils is not adequate in obtaining data that has high reliability. In general, the Handball Cognitive Assessment Instrument has a reliability of 0.65 and 0.68 for the primary test and repeated test which it is considered the value is sufficient to be used in the study (Sumintono & Widhiarso, 2015).

DISCUSSION

The Handball Cognitive Assessment instrument is a brief cognitive assessment test that was specifically developed to evaluate the knowledge and understanding of students based on the cognitive domain for the category of handball attack-play. Several studies previously by Anderson et al., (2001); Darst & Pangrazi, (2009); Chang et al., (2020), has a supportive finding that the cognitive domain is a vital component in the acquisition of PE since it contributes to the successful learning of physical activities. Thus, the general goals of this study were to create a cognitive evaluation tool for Year 4 primary school students learning handball, as well as to confirm the instrument's validity and reliability using Rasch Model analysis.

The analysis focused on identifying evidence for construct validity for dichotomous data, which was critical in determining how far the item measure could accurately reflect what needed to be measured based on the ideas and criteria established (Nurfaizin, 2019). The subsequent analysis is the unidimensionality test for the structure of the developed instrument by referring to the Principal Component Analysis of Rasch Residuals to find out the validity of the instrument (Andrich, 1988). As for the reliability, individual isolation indexes reveal a substantial amount of competence in a set of sample respondents, whereas item isolation indexes show how far the sample respondents can display item difficulty isolation (Boone & Noltemeyer, 2017). Furthermore, using an instrument as a dependable data gathering tool that can also be depended on to reveal the genuine conditions and situation on the ground is vital (Fahruna and Fahmi 2017).

To determine the construct validity of an instrument, a theoretical evaluation of the notion of the variable to be measured must be conducted, beginning with the formulation of the construct, determining dimensions and indicators, and ending with the development and drafting of the instrument's items (Erfan et al., 2020). The construct must be developed based on a rational and thorough research and comparison process that synthesizes hypotheses about the concept of the variable to be measured. The construct validation process of an instrument must be conducted through expert review or justification or through the assessment of a set of panels consisting of people who grasp the topic or content of the variables to be assessed, as indicated in the theoretical review process (Ariffin et al., 2010).

The first analysis carried out on the items was an analysis of construct validity which was done by looking at polarity items. The results showed the polarity of the items of the entire construct was in the range of 0.01 to 0.51. Following the criteria recommended by Linacre, (2009), the Point Measure Correlation (PT-MEASURE CORR) or PMC value must be positive, and all the item is appropriate. This

shows that there is no conflict between item problems with the measured question construction. Besides, as the logit value is found to be nearest to 1.5, the researchers concluded that keeping all the cognitive instrument's items is appropriate (Erfan et al., 2020).

In subsequent analysis, researchers use the Principal Component Analysis (PCA) of Rasch Residuals to determine the unidimensional construct of the Handball Cognitive Assessment tool. Unidimensionality assumes that the things being measured can only measure a single effort. This model outlines a simple process for determining a measurement's dimension. One potential threat to unidimensionality, according to Rasch, is a side factor. According to Baghaei and Aryadoust, (2015), the value of unexplained variance in the first construct residual PCA greater than 15% is poor, 10-15% is adequate, 5-10% is strong, 3-5% is very strong, and less than 3% is exceptional. The value of unexplained variance in the first construct is at 8,4% indicating sufficient and strong indicators. The value of the variance size (Variance explained by measures) is best > 40%. However, it shows a value of 25.4% which was the minimum requirement of 20% in this study (Sumintono & Widhiarso, 2015). This is further explained by Ramdani et al., (2020), if the measurements can explain the raw variation to a ≥ 20% level, it shows that unidimensional measurement can be justified. The interpretation criteria are as follows: satisfactory if the percentage is between 20% and 40%, good if the percentage is between 40% and 60%, and unbelievably good if the percentage is greater than 60%.

A repeated test with a gap of two-week from the primary test reveals almost the same result which indicates the instrument has good reliability overall. Accordingly, Linacre (2009) recommends classifying items and individuals with a reliability rating of 0.67 to 0.80 as simple, 0.81 to 0.90 as good, and > 0.91 as very good. Meanwhile, the reliability index of items and respondents received was 0.83 and 0.54 respectively for both test and repeated tests. The Cronbach Alpha value (KR-20) shows a value of 0.65 for the primary test and 0.68 for the repeated test with a slight gain showing the reliability of the instrument is adequate (Sumintono & Widhiarso, 2015). In addition, the Cronbach Alpha value (KR-20) means if the items were assessed using classical test theory, this value indicates that the results of a satisfactory or consistent reliability study were obtained (Erfan et al., 2020). As details of the reliability analysis, the consistency of the respondents' responses is low, while the quality of the instrument items is good.

Overall, the validity and reliability of the Handball Cognitive Assessment test are good. As a result, this instrument is adequate and can be used to assess a student's knowledge and understanding of the category of handball attack-play based on the cognitive domain. However, this study has two drawbacks, according to the researcher's knowledge. To begin, the data for this study was gathered utilizing a Google form using a quiz approach in the form of an assessment, in which the student might seek assistance from their parents. As a result, the data's quality may be a source of worry. A prospective study may be conducted in the future to cross-validate our findings. Second, our data came from a single school in the Kuala Lumpur zone, which limits applicability and comparability.

CONCLUSION

The conclusion that can be made from this study is that based on the results of the test of construct validity which consists of 25 items, it was found that 22 out of the 25 items tested do not contradict with any of the items, and the development of the item that was measured with the value of Outfit-MNSQ that is less or equal to the value of logit 1.5. For items 6, 8, and 14, although the Outfit-MNSQ exceeds logit 1.5, when it is compared to the primary test, it did not diverge too far from the logit value. Furthermore, from the item reliability test, it is found that the reliability value of items/questions is 0.83 and the index of individual respondents is 0.54. Hence, the Handball Cognitive Assessment Instrument is dependable and can be used to measure the cognitive domain for knowledge and understanding of the topic of Handball in primary Year 4 students. Thus, the study needs to be expanded to other chapters of the PE syllabus following the concepts outlined. This is to ensure assessment is done using valid and reliable developed instrument. Besides, this study appears to be a step stone for reference and motivating resource for PE teacher committee with the

goal of improving PE teaching and learning in future by having holistic assessment for the student involving the cognitive domain in PE which seen to left out previously.

REFERENCES

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association : JMLA*, *103*(3), 152–153. https://doi.org/10.3163/1536-5050.103.3.010

Akın, A. (2022). The effectiveness of web-based Mathematics instruction (WBMI) on K-16 students' mathematics learning: a meta-analytic research. *Education and Information Technologies*. https://doi.org/10.1007/s10639-022-10931-x

Amer, A. (2006). Reflections on Bloom's revised taxonomy. *Electronic Journal of Research in Educational Psychology*, *4*(1), 213–230.

Anderson, L. W., Krathwohl Peter W Airasian, D. R., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *Taxonomy for_ Assessing a Revision of Bloom's Taxonomy of Educational Objectives*. https://www.uky.edu/~rsand1/china2018/texts/Anderson-Krathwohl - A taxonomy for learning teaching and assessing.pdf

Ayers, S. F. (2004). High school students' physical education conceptual knowledge. *Research Quarterly for Exercise and Sport*, *75*(3), 272–287. https://doi.org/10.1080/02701367.2004.10609160

Baghaei, P., & Aryadoust, V. (2015). Modeling Local Item Dependence Due to Common Test Format With a Multidimensional Rasch Model. *International Journal of Testing*, *15*(1), 71–87. https://doi.org/10.1080/15305058.2014.941108

Balakrishnan, M., Rengasamy, S., & Aman, M. S. (2011). *Teaching Game for Understanding in Physical Education : A Theoretical Framework and Implication*. *1*(2), 201–214.

Bhasah, & Bakar, A. (2008). *Pengujian, pengukuran dan penilaian pendidikan.* Prospecta Printers Sdn. Bhd.

Bloom, B. S., Englehart, N. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy Of Educational Objectives: The Classification Of Educational Goals, Handbook I: Cognitive Domain.* David McKay Company.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model*. Routledge. https://doi.org/10.4324/9781315814698

Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education*, *4*(1), 1416898. https://doi.org/10.1080/2331186X.2017.1416898

Cannon, H. M., & Feinstein, A. H. (2005). Bloom beyond Bloom: Using the revised taxonomy to develop experiential learning strategies. *Developments in Business Simulations and Experiential Exercises*, *32*, 348–356.

Chang, K. E., Zhang, J., Huang, Y. S., Liu, T. C., & Sung, Y. T. (2020). Applying augmented reality in physical education on motor skills learning. *Interactive Learning Environments*, *28*(6), 685–697. https://doi.org/10.1080/10494820.2019.1636073

Darst, P. W., & Pangrazi, R. P. (2009). *Dynamic physical education for secondary school students*. San Francisco : Pearson/Benjamin Cummings, 2009. http://hdl.handle.net/11162/65421

Fahruna, Y., & Fahmi, M. (2017). Validitas dan Reliabilitas Konstruk Pengukuran Perpustakaan Ideal Berbasis Pemakai dengan Pendekatan LIBQUAL. *Jurnal Ekonomi Bisnis Dan Kewirausahaan*, *6*(2), 161. https://doi.org/10.26418/jebik.v6i2.22989

Ferris, T., & Aziz, S. (2005). A psychomotor skills extension to Bloom's taxonomy of education objectives for engineering education. *Exploring Innovation in Education and Research*, 1–7.

Gengatharan, K., & Azli Rahmat. (2019). Keperluan modul pentaksiran pendidikan kesihatan untuk guru tahap satu dalam pelaksanaan pentaksiran bilik darjah. *Jurnal Sains Sukan Dan Pendidikan Jasmani 8(2): 24-34*, *8*(2), 19–27.

Gurvitch, R., & Metzler, M. (2013). Aligning Learning Activities with Instructional Models. *Journal of Physical Education, Recreation & Dance*, *84*(3), 30–37. https://doi.org/10.1080/07303084.2013.767719

Kamarudin, H., & Siti Hajar, A. A. (2004). *Pedagogi asas pendidikan. :* Kayazano Enterprise.

Linacre. (2009). *Winsteps (Version 3.68) [Computer Software]*. Beaverton, Oregon: Winsteps.com

Liza, S., Julismah, J., Azali, R., & Norkhalid, S. (2016). Learning Assessment Model for Invasion Category Games in Year 4 Physical Education. *Journal of Sports Science and Physical Education*, *5*(1), 35–48.

Misbach, I. H., & Sumintono, B. (2014). Pengembangan dan validasi instrumen "persepsi siswa tehadap karakter moral guru" di Indonesia dengan model rasch. ,. *PROCEEDING Seminar Nasional Psikometri*, 148–162.

Mohan, D. (2018). Flipped Classroom, Flipped Teaching and Flipped Learning in the Foreign/Second Language Post–Secondary Classroom. *Nouvelle Revue Synergies Canada*, *11*(11), 1–12. https://doi.org/10.21083/nrsc.v0i11.4016

Nixon, J. E., & Locke, L. (1973). Research in teaching in physical education. In *(in:) R.M.W. Travers, ed., Second handbook of research on teaching* (pp. 1210–1242). Chicago, IL: Rand McNally.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, *30*(4), 459–467. https://doi.org/10.1002/nur.20199

Ramdani, R., Hanurawan, F., Ramli, M., Lasan, B. B., & Afdal, A. (2020). Development and Validation of Indonesian Academic Resilience Scale Using Rasch Models. International Journal of Instruction, 14(1). https://doi.org/10.29333/IJI.2021.1417A

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan.* Trim Komunikata Publishing House.

Sumintono, Bambang, & Widhiarso, W. (2014). *Aplikasi Model Rasch untuk Penelitian Ilmu-Ilmu Sosial (edisi revisi)*.

Zhang, T., Chen, A., & Ennis, C. (2019). Elementary school students' naïve conceptions and misconceptions about energy in physical education context. *Sport, Education and Society*, *24*(1), 25–37. https://doi.org/10.1080/13573322.2017.1292234

Uthaya Chanthira Kumar a/l Muniandy
Department of Mathematics & Science Education
Faculty of Education
University Malaya
Kuala Lumpur, Malaysia.
Email: uthaya8c5kumar@yahoo.com